Emotion-Inducing Imagery versus Motor Imagery for a Brain-Computer Interface

Emotion-Inducing Imagery versus Motor Imagery for a Brain-Computer Interface

A. D. Bigirimana, N. Siddique, *Senior Member, IEEE*, and D. Coyle, *Senior Member, IEEE*

***Abstract*— Neural correlates of intentionally induced human emotions may offer alternative imagery strategies to control brain-computer interface (BCI) applications. In this paper, a novel BCI control strategy i.e., imagining fictional or recalling mnemonic sad and happy events, emotion-inducing imagery (EII), is compared to motor imagery (MI) in a study involving multiple sessions using a two-class electroencephalogram (EEG)-based BCI paradigm with 12 participants. The BCI setup enabled online continuous visual feedback presentation in a game involving one-dimensional control of a game character. MI and EII are compared across different signal-processing frameworks which are based on neural-time-series-prediction-preprocessing (NTSPP), filter bank common spatial patterns (FBCSP) and hemispheric asymmetry (ASYM). Online single-trial classification accuracies (CA) results indicate that MI performance across all participants is 77.54% compared to EII performance of 68.78% ($p<0.05$). The results show that an ensemble of the NTSPP, FBCSP and ASYM frameworks maximizes performance for EII with average CA of 71.64% across all participants. Furthermore, the participants' subjective responses indicate that they preferred MI over emotion-inducing imagery (EII) in controlling the game character, and MI was perceived to offer most control over the game character. The results suggest that EII is not a viable alternative to MI for the majority of participants in this study but may be an alternative imagery for a subset of BCI users based on acceptable EII performance (CA > 70%) observed for some participants.**

***Index Terms*— BCI, EEG, emotion-inducing imagery, motor imagery, games, neurogaming, assistive technology, machine learning, AI.**

I. Introduction

BRAIN-computer interfaces (BCIs) offer a means to communicate and control computer-based applications without movement, including entertainment [1] [2] (e.g. BCI games), vehicle control [3], rehabilitation [4] and assistive technologies. BCIs are built around decoding the person's intent by direct measurement of brain activity [5], usually measured through electroencephalography (EEG). One of the challenges in BCI is that there are limited control strategy options available to users: some strategies, e.g., motor imagery, are challenging for some users and require training [6][7], and other strategies (evoked potentials) often require gaze control and are dependent on external stimuli. As a non-negligible portion of users are unable to learn to control a motor imagery (MI)-based BCI [6] within a limited duration of training, and there is ongoing debate on how such users should be characterized in BCI research community [8]. There is a necessity to investigate alternative imagery strategies to avail more options for BCI users. Different alternative imagery strategies that have been used to control BCIs include emotion inducing imagery (EII) [9] [10] [11] [12], mental singing [13], mental arithmetic [14], mental rotation, word association, auditory imagery, mental subtraction, and spatial navigation [15]. In some cases, alternative imagery strategies outperformed motor imagery. Curran and colleagues reported better ease of use and higher classification performance with spatial navigation and auditory imagery compared to motor imagery [16]. In this paper, the viability of EII as a potential BCI control strategy is investigated. We are interested in assessing EII strategy because the associated tasks can be constructed from the user's natural experience.

The EII strategy exploits the differences observed in brain responses to different emotional stimuli or recall of emotional events, and this may even enable a multi-class BCI [9]. Positive emotions (e.g., happy, joy) are associated with less relative alpha power in left frontal cortical regions than the right, whereas for negative emotions (e.g., sad, disgust) less relative alpha power is observed in the right frontal cortical area [17] [18], and similar hemispheric asymmetry activation was reported in functional imaging [19]. In addition to the differences in brain activity associated with different emotions, for emotion to be useful in active, independent BCIs, where the user issues a command as opposed to waiting on a stimulus to evoke a brain response, the BCI user is required to imagine or recall emotional situations. Kothe et al. [20] reported an accuracy of 71.3% across participants in a two-class classification of valence ratings, where the participants were self-paced in recalling positive and negative valence emotions. In a similar study, Chanel et al [10] achieved an accuracy of 63% in a three-class (negative emotion, positive emotion, and neutral) and 80% for two-class classification, where each emotion-inducing imagery task lasted about 8s. Furthermore, Iacoviello and colleagues [11] achieved a classification accuracy of 90.2% for imagery induced by remembering an unpleasing odor versus a relaxed state. Sitaram and colleagues [12], in an fMRI-based study, presented performance-based feedback to participants who were recalling sad, happy, and

This research is supported by a Vice Chancellor's Research Scholarship at Ulster University.

A. D. Bigirimana, N. Siddique, and D. Coyle are with Intelligent Systems Research Centre, Ulster University, Derry, BT48 7JL, UK (e-mails: Bigirimana-a@ ulster.ac.uk, nh.siddique@ulster.ac.uk, dh.coyle@ulster.ac.uk)

disgust emotions (three-classes), and achieved an accuracy of 60% classification with feedback presentation.

Only a limited number of previous studies have applied emotion-inducing imagery with real- or pseudo-real-time feedback presentation. In a typical BCI system, the user should be provided with feedback interaction as the feedback implicitly acts as a motivating reward to the user during brain modulation tasks [21] and helps users learn and develop their imagery strategy. In the preliminary EII studies [22][23], participants controlled a video game character by recalling sad and happy events, and their performance was not significantly different from performance achieved using classical right versus left hand motor imagery. However, further investigation across multiple sessions is necessary to establish the viability of EII compared to motor imagery.

In the current study, we present a comprehensive analysis of the viability of EII, comparing multiple online feedback sessions and single-trial classification accuracy for EII tasks versus MI with 12 participants. Since the participants participated in multiple sessions, we also investigated the potential correlation between their online performance and the relative band power in various frequency bands of EEG recorded prior to each run i.e., before the subjects engage in MI and EII tasks. We also investigated the most relevant frequency bands and scalp areas for MI and EII. The participants' online performance across multiple sessions with EII and MI during one-dimensional control of a video game character are reported. As the neural correlates of EII differ from that of MI, state-of-the-art methods for MI classification may not suit EII, we therefore compare the performance of MI and EII using a range of signal processing frameworks: neural-time-series-prediction-preprocessing (NTSPP) framework, a framework based on the hemispheric asymmetry (ASYM), and framework based on filter-bank common spatial pattern (FBCSP). In addition to EEG based analysis, the participants' subjective responses to a questionnaire are reported. This questionnaire assessed participants' favorite imagery strategy as a BCI control approach and which imagery they perceived to offer most accurate feedback.

II. Material and Methods

A. Participants

The Study involved 12 healthy volunteering participants (2 females and 10 males, mean = 29 years, SD = 8), recruited at Ulster University. All these participants had normal or corrected to normal vision and were not suffering from any condition that would impede their participation in BCI experiments. The participants reported no neurological disease or mental illnesses. The study was approved by Ulster University's Faculty of Computing, Engineering and Built Environment research ethics committee, and written informed consent was provided by each participant prior to participating in the study. Two participants had previously participated in BCI studies and had good performance in MI. Ten participants had no previous BCI participation experience prior to this study. Participants were given some practice and demonstration to get them comfortable with the experiment before participating in the first session.

B. Experiment Setup

This study was organized in multiple sessions scheduled on different days. Each session includes four runs: two EII runs and two MI runs. In each session there was always a minimum of two feedback runs i.e., feedback of each imagery type was provided for at least one run in every session as shown in Fig. 1. In each session, the participant begins with either EII or MI in the first run (selected randomly), and for the remaining runs of that session, he/she alternates between EII and MI i.e., there were no successive EII runs or MI runs in one session. There were rest breaks of 2 to 5 minutes between runs. Each run started with participant relaxing while minimizing eye-blinks for a period of 60 s followed by 60 trials, randomly ordered, 30 trials per class. The session format and trial timing are shown in Fig. 1.

1) Imagery Tasks, Cue, and Feedback Presentation

Before starting the sessions, the participant was instructed to identify two real or fictitious emotional events: one event that the participant considers as a sad event and another event that he/she considers as a happy event. To avoid possible emotional stress in the participants, they were instructed to refrain from using extremely sad events. Participants were instructed to keep the event very brief and preferably to focus on the most emotional episode of the event. The tasks during EII run are to recall the happy event when the presented cue appears on the right-hand side on the computer screen and to recall the sad event if the cue appears on the left side. The tasks for MI run are to imagine moving the right hand when the cue is on the right and to imagine moving the left hand if the cue appears on the left side (without actually moving the hands). Participants were

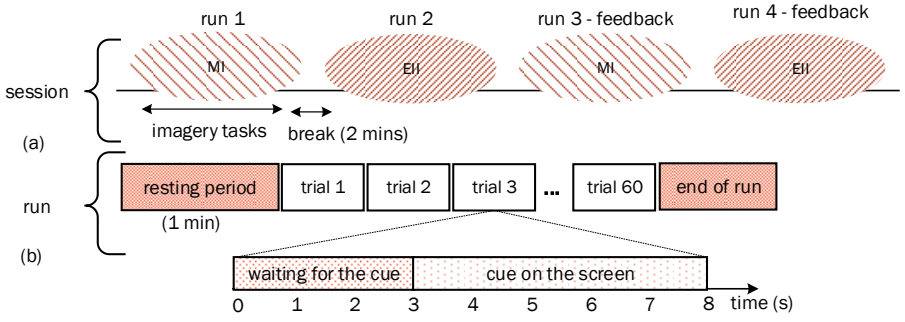

Fig. 1. (a) The recording session structure and (b) run structure with the structure of each trial in the run.

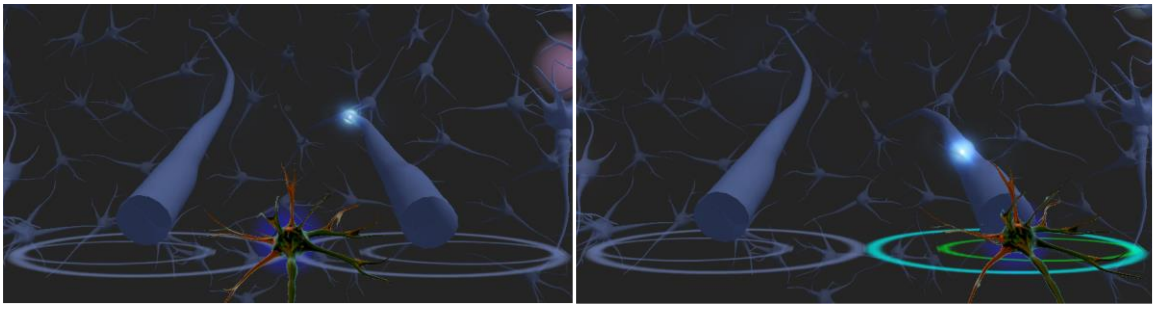

Fig. 2. The screen shots for cue and feedback presentation using the Neurosensi Games Platform

instructed to be consistent in their imagery strategy during the session.

In each run, we utilized a computer game paradigm called NeuroSensi to cue the tasks. In this game, a light, representing a neuronal action potential (or spike), travels through one of the two graphical axons (left-side or right-side axon) on the computer screen, see Fig. 2. The appearance of the spike (light) cues the imagery task. Once the spike reaches the end of the axon and disappears, the participant stops the imagery task, relaxes and waits for the next spike. The time on and off the screen of the cueing spike follows the trial structure (0 to 8s of the shown segment of a run) in Fig. 1, where the cue appears at 3s and disappears at 8s.

In the feedback run, a continuous feedback is given as a horizontally moving game character (a graphical representation of neuron's cell body and dendrites), and the game challenge is to move this character to collect the light (spike), shown on the right in Fig. 2. Points are awarded for moving the game character in the right direction and positioning the character as close as possible to the axon when the spike reaches the end of the axon. Additional points are awarded for collecting more than three spikes consecutively without failure. These bonus points are accompanied with background neurons firing and propagating several spikes, extending the waiting period (after the cueing spike disappears) by 2s. This continuous feedback, i.e., movement of the game character, is controlled by the BCI. Participants were instructed to focus on executing the cued task during the task execution as much as possible. This instruction was given to reduce potential frustration resulting from poor classification of the task.

2) Participants' Subjective Responses

After each experiment session, each participant reported his/her favorite control approach and the imagery strategy perceived to provide best control over the game character. This was done through a short questionnaire where the participant answered by selecting "*MI",* "*EII"*, or "*equally the same"* on the questionnaire.

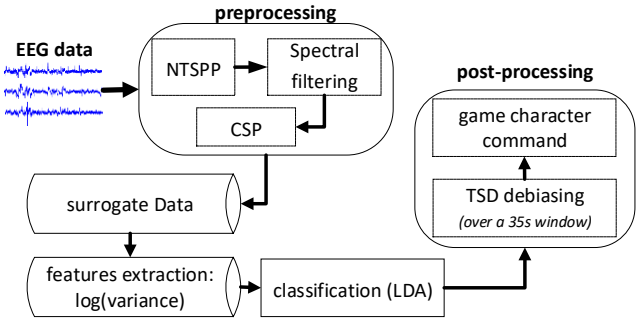

Fig. 3. The NTSPP framework

C. Data Analysis

1) EEG Data

We acquired EEG data using g.tec (Guger Technologies, http://www.gtec.at/Products) biosignal amplifiers (g.BSAmp) setup with 30 active EEG electrodes (g.GAMMAsys, g.Ladybird) positioned in a 10-20 system plus two electrooculagram electrodes (F3, F4, FC5, FC1, FC2, FC6, C3, CZ, C4, CP5, CP1, CP2, CP6, P3, PZ, P4, AF3, AF4, F7, FZ, F8, T7, T8, P7, P8, PO3, PO4, O1, OZ, O2, HEOG, and VEOG). The data were recorded with sampling rate of 250Hz then down-sampled to 125Hz. We visually inspected the data recorded in non-feedback runs for significant artefacts (e.g., eye-blinks), and then carried out an offline analysis to find optimal parameters to be applied in feedback runs.

2) Signal Processing Frameworks

We compared the performance of EII and MI across different signal processing frameworks: neural-time-series-prediction-preprocessing (NTSPP) framework [24][25], a hemispheric asymmetry (ASYM) [17] framework, a filter-bank common spatial pattern (FBCSP) framework [26], framework combining NTSPP and FBCSP (NTSPP-FBCSP), and framework combining NTSPP, FBCSP and ASYM (referred to as COMB in this paper). NTSPP is a framework that has been extensively at our BCI lab with MI based BCI. FBCSP is one of the state-of-the-art frameworks for MI based BCI [26]. ASYM framework is inspired by the hemispheric asymmetry of emotions; different asymmetry features are usually used for emotion recognition [27].

In each of these frameworks, the data for the no-feedback run are used to generate optimized settings which are deployed to drive the feedback in the online run or to simulate online run.

a) NTSPP Framework

The NTSPP framework produces a surrogate data space which is more separable through neural network based specialization in time-series prediction of individual EEG channels [24] [25] combined with spectral filtering (SF) in 8 – 30Hz frequency band and common spatial patterns (CSP) to

maximize the separability between classes [28] [29]. The NTSPP framework is illustrated in Fig. 3.

In the NTSPP different prediction networks are trained to specialize in predicting future samples of EEG signals on each channel. Due to network specialization, features extracted from the predicted signals are more separable and thus easier to classify. The number of time-series available and the number of classes governs the number of specialized predictor networks and the resultant number of predicted time-series from which to extract features.

$$P = M \times C, \tag{1}$$

where P is the number of networks (which is equal to number of predicted time-series), M is the number of EEG channels and C is the number of classes. The prediction follows the expression in (2),

$$\hat{x}_{ci}(t+\pi) = f_{ci}\langle x_i(t), \ldots, x_i(t-(\Delta-1)\tau\rangle, \tag{2}$$

where t is the current time instant, Δ is the embedding dimension and τ is the time delay, π is the prediction horizon, f_{ci} is the prediction model trained on the i^{th} EEG channel, x_i, i=1,..,M, for class c, c =1,..C, and $\hat{x}_{ci}$ is the predicted time-series produced for channel i by the predictor for class c. Each prediction network, f_{ci} in this work is a self-organizing fuzzy neural network (SOFNN) [30].

Prior to the calculation of the spatial filters, X is spectrally filtered in a specific frequency band, 8 – 30Hz. This band encompasses the alpha, beta bands which are relevant during sensorimotor processing [30] [31], and these bands or sub-bands within these bands are often used for emotional states detection [32]. CSP is used to maximize the ratio of class-conditional variances of EEG sources. CSP is applied by pooled estimates of the covariance matrices, Σ_1 and Σ_2, for two classes, as follows:

$$\sum_c = \frac{1}{I_c}\sum_{i=1}^{I_c} X_i X_i^t \; (c \in \{1,2\}), \tag{3}$$

where I_c is the number of trials for class c, and X_i are the $M\times N$ matrices containing the i^{th} windowed segment of trial i; N is the window length, and M is the number of EEG channels – when CSP is used in conjunction with NTSPP, $M = P$ as per (1). The two covariance matrices, Σ_1 and Σ_2, are simultaneously diagonalized such that the Eigenvalues sum to 1. This is achieved by calculating the generalized eigenvectors W:

$$\sum_1 W = (\sum_1 + \sum_2) W D, \tag{4}$$

where the diagonal matrix D contains the eigenvalues of Σ_1, and the column vectors of W are the filters for the CSP projections. With this projection matrix the decomposition mapping, E, of the windowed trials X is given as:

$$E = WX. \tag{5}$$

Features, $\bar{\omega}$, are derived from the log-variance of preprocessed/surrogate signals, E, within a 2s sliding window:

$$\bar{\omega} = log(\, var(\, E)). \tag{6}$$

The dimensionality of $\bar{\omega}$ depends on the number of surrogate signals used from E. The common practice is to use several (between 2 and 6) eigenvectors from both ends of the eigenvector spectrum i.e., the columns of W. Using NTSPP the dimensionality of X can increase significantly. CSP, can be used to reduce the dimensionality therefore combining NTSPP with CSP leads to increased separability while maintaining a tractable dimensionality [25]. The number of CSP filters and time points with maximum separability are assessed using leave-2 trials (one trial from each class)-out cross-validation on a 2s sliding window features with linear discriminant analysis (LDA).

The optimized parameters, number of CSP filters and identified time point of maximum separation, are used to setup the final classifier using all the training data. This classifier is then deployed online, in a MATLAB© Simulink model. In the online processing, the classifier's output translation to the game character movement is de-biased to account for class bias behavior and to improve feedback stability by continuously removing the mean of recent classifier outputs. This mean is computed from a 35s window on the most recent classifier outputs. At each sample point, the classifier's output is a distance computed from the classifier's learned weights vector; often referred to as time-varying signed distance (TSD) [28][33]. The TSD value at a given time point t during n^{th} trial is given by expression in (7). The distance's sign indicates the classifier's output label and its magnitude measures the classification confidence. The magnitude of the TSD indicates how far the game character moves, and the sign indicates the direction of the character's movement (moving to the right or to the left). The value of current TSD is de-biased by subtracting from it the mean of TSD values for the previous 35s. After the feedback run, the continuous classification performance is assessed by computing the percentage of trials having TSD (de-biased) values with the same sign as the targeted class, at each time point t.

$$TSD_t^{(n)} = w^T \bar{\omega}_t^{(n)} - a_0 \tag{7}$$

where w^T and a_0 are slope and bias of the discriminant hyperplane, respectively, of the trained LDA trained, $\bar{\omega}_t^{(n)}$ is the features vector at the time point t of the nth trial.

b) Hemispheric Asymmetry (ASYM) Framework

Electrodes located on the left and right hemispheres, at equivalent positions, are paired into thirteen pairs for asymmetry feature extraction in different frequency bands: theta (4-7Hz), alpha (8-12Hz), beta (13-30Hz), and low gamma (31-45Hz). Differential, D, and ratio, R, asymmetry features are extracted from a 2 s sliding window.

$$D_{asym} = r_i - l_j \tag{8}$$

$$R_{asym} = r_i / l_j \tag{9}$$

where the r_i is band power in the frequency band i at the right-side electrode and l_j the band power in the frequency band j at the left-side electrode of the pair.

The parameters to be optimized in this framework are the number of the features and the best time in the trial for classifier

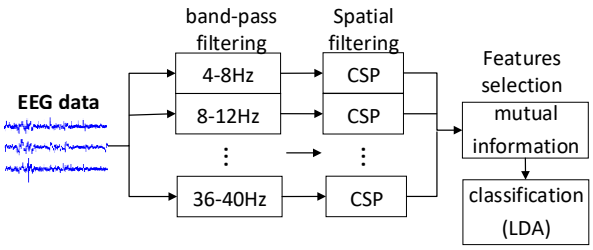

Fig. 4. FBCSP framework setup

training. The parameters are optimized through a leave-2 trials (one trial from each class)-out cross-validation setup. The final number of features varies from 4 to 16. These features are selected based on *mutual information best individual feature* (MIBIF) algorithm [34]. The MIBIF computes mutual information for each individual feature, and a subset of features with highest mutual information is selected. An LDA classifier is trained with optimized parameters and then deployed in a MATLAB© Simulink model that simulates an online run with data from the feedback run.

c) Filter Bank Common Spatial Patterns (FBCSP) Framework

In the FBCSP, the EEG data from the calibration run (run without feedback) are filtered in nine different frequency bands (4-8Hz, 8-12Hz, 12-16Hz, 16-20Hz, 20-24Hz, 24-28Hz, 28-32Hz, 32-36Hz, and 36-40Hz), and then CSP-based features, as in (6), are extracted from each band-filtered signal as shown in the Fig. 4. The parameters considered in this setup are the number of features, number of CSP filter pairs, and the best time during the trial to train the classifier.

The CSP features (log-variance) are extracted on a 2s sliding window, and the number of CSP filter pairs together with number of features (4 to 16 features) are optimized in a 6-fold-cross-validation setup with an inner 5-fold-cross-validation. The best features are selected based on MIBIF algorithm. The peak for the average time-course given by six time-courses of accuracy from the six folds is used as the best time (during the task execution) period for training the classifier. After optimizing the parameters, an LDA classifier is trained and then deployed in a MATLAB© Simulink model that simulates the online run with data from feedback run.

d) NTSPP-FBCSP and COMB Frameworks

We also evaluated the CA performance with signal processing frameworks that combine different setups. We combined the outputs from individual frameworks into one output during online re-simulation. Before combining the frameworks' outputs, each framework's output is individually de-biased as in (7), and the combined output is de-biased in the same way. NTSPP was combined with FBCSP, and this was motivated by the fact they are both CSP-based. Additional motivation was to augment the NTSPP-based setup that uses one wide frequency band, by the FBCSP-setup which utilizes nine narrower frequency bands. The output TSD of NTSPP-FBCSP combination was given by the individual framework's TSD with highest magnitude among the two as shown in (10).

$$TSD_{NTSPP-FBCSP} = s \times max(|TSD_{NTSPP}|, |TSD_{FBCSP}|) \quad (10)$$

where s is the sign of the TSD with the highest absolute value.

Apart from NTSPP-FBCSP, the initial three frameworks (NTSPP, ASYM, and FBCSP) were combined into one framework, COMB. For the COMB framework, the output TSD is given by the mean of the individual TSDs which have matching signs, i.e. if two or three of the three TSDs match their signs, the overall TSD for the combination is the average of those individual TSDs.

3) Contribution of Different Electrodes and Frequency Bands

In the order to identify the most relevant frequency bands and scalp areas for MI and EII during the task execution, we run the feedback runs' data, as training data, through FBCSP framework in an offline analysis. After optimizing the parameters (number of CSP filters pairs, number of features, and time-point for the peak cross-validation accuracy) through cross-validation as previously described, a final CSP projection matrix W (as in (4)) is computed for each frequency band. Each projection matrix W, gives us the weights for each electrode's contribution to the surrogate data resulting from CSP filtering in each frequency band. Log-variance features are then extracted with a 2s sliding window. The global weights for electrodes at a time-sample t, are given by the weights matrix $K(t)$, computed in (11), from projection matrices across nine frequency bands and mutual information associated to extracted features using the projection matrices at a given time-sample. To average topographic $K(t)$ (CSP-MIBIF) weights, across different sessions at a given time-sample, topographic CSP-MIBIF weights are first normalized to the electrode with highest weight for each individual run.

$$K(t) = \sum_{n=1}^{9} W_n \cdot Y_n(t) \quad (11)$$

where W_n is the projection matrix in the frequency band n, and $Y_n(t)$ is the mutual information (weights) given to the features extracted with W_n at the time-sample t in frequency band n.

Apart from the CSP-MIBIF weights of electrodes, the CSP-MIBIF weight of each frequency band is computed by adding the mutual information of features extracted from the same frequency band. At each time-sample (from 2s in the trial as we use a 2s window), we extract features with optimized parameters and then compute the mutual information for each feature using MIBIF. This allows us to establish a time-course of weights for the nine frequency bands for each feedback run and to compute an average time-course of weights from several feedback runs.

4) Pre-run EEG Analysis

Resting state EEG recorded at the beginning of each run was investigated retrospectively to determine if the spectral power ratios for specific frequency bands, with respect to total frequency content, may be used as a predictor for performance during BCI. Previously, high ratio for theta and low ratio for alpha were reported to be associated with poor performance in motor imagery [7]. In the present study, we compute the spectral power ratios for theta, alpha, beta, and gamma bands at

five topographical areas covered by EEG electrodes: frontal, temporal, central, parietal and occipital area.

Before computing power ratios for various frequency bands, the channels with noise were removed from the data by applying 'kurtosis' and 'spectrum threshold', utilities of the EEGLAB toolbox [35]. The data were then filtered with a high-pass filter (0.5Hz) followed by further automated artefact removal using a hybrid independent component analysis (ICA) – wavelet transform (WT) [36] [37]. In this ICA-WT analysis, the *runICA* algorithm from the EEGLAB toolbox was applied on the data, and the resulting independent components were individually decomposed in wavelet coefficients by wavelet transform. The wavelet coefficients were thresholded then followed by reconstructing independent components by an inverse wavelet transform. The reconstructed independent components were re-mixed to produce clean EEG data.

The cleaned data were filtered by a low-pass filter (45Hz), and the ratio, *R*, for each band was calculated as the spectral power of the signal filtered in a given frequency band divided by the total spectral power in the signal. The spectral power was computed as in (12):

$$p = (1/m)\sum_{t=1}^{m} x(t)^2, \tag{12}$$

where *x* is the band-pass filtered signal, and *x(t)* is the signal sample at the time *t*, with *t = 1, 2, 3… m*; *m* being the number of samples in the signal *x*.

$$R_b = p_b/p_{total} \tag{13}$$

where *b* is one of the frequency bands, and *total* is the entire 0.5 – 45Hz band.

We computed Pearson correlation between each frequency band's ratios and online classification accuracies achieved across different sessions for each participant.

D. Statistical Tests

In this study we compare single-trial classification accuracy (CA) performance for two imagery approaches (MI and EII) across different signal processing frameworks (FBCSP, ASYM, NTSPP-FBCSP, and COMB). Wilcoxon signed rank test, with significance level of 0.05, is used to evaluate the difference between online CA of MI and online CA of EII across the participants during the feedback runs. Furthermore, repeated analysis of variance (ANOVA), with a significance level of 0.05, is used to compare the two imagery approaches' performance across the five frameworks considered in this study, i.e. NTSPP (used in the actual online setup) and the frameworks used in the re-simulation of feedback runs.

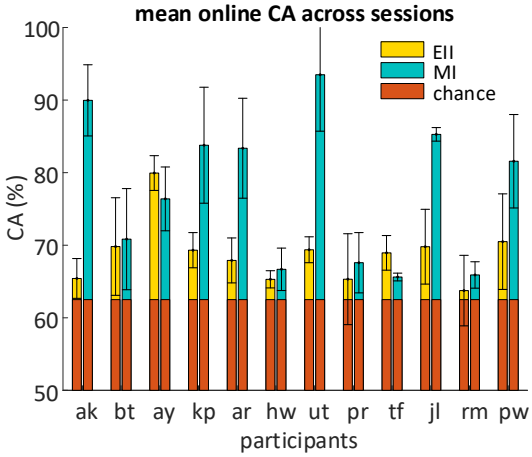

Fig. 5. Online single trial classification accuracies averaged across sessions for each participant with theoretical chancel level (random-CA)

III. Results

A. Classification Accuracy (CA)

The online single-trial CA for feedback runs averaged across sessions for each participant are reported in Fig. 5. All the participants performed above chance level in all their MI sessions. For EII, on the other hand, only 9 participants performed above chance level in all their sessions. The participant *'ak'*, *'pr'*, and '*rm*' each performed below chance level in one of their EII sessions (the chance level upper limit is 62.39% for a 2-class problem, 30 trials per class, with 95% of confidence interval [38]), but the average performance across their sessions is above the chance level for each of these three participants. Wilcoxon signed rank test showed that averaged online MI CA are higher than EII CA ($p < 0.05$).

The re-simulated online single-trial classification accuracy results are reported in *Fig. 6* (NTSPP results are from actual online runs with participants). The framework with highest averaged CA across participants, for EII, is the COMB framework with averaged CA of 71.64% across participants. The frameworks with lowest performance in EII case are observed in FBCSP and ASYM frameworks with average CAs across participants of 66.82% and 66.92%, respectively. For MI, the best performing framework is the NTSPP-based framework with average CA of 77.54% across participants, and ASYM-based framework least performing with average CA of 67.05%.

The repeated analysis of variance (ANOVA) results show

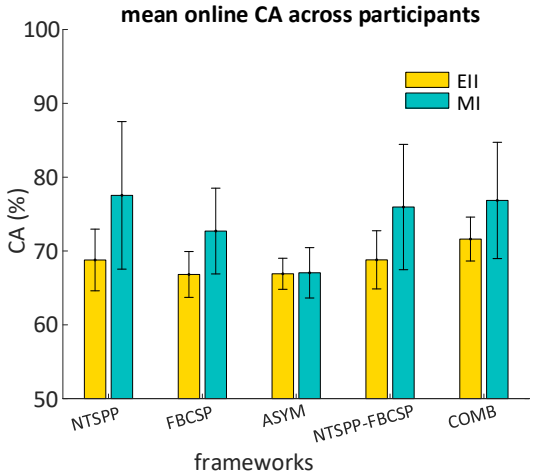

Fig. 6. The online single-trial CA during recording (with NTSPP) and re-simulated single-trial CA (with FBCSP, ASYM, NTSPP-FBCSP, and COMB frameworks) averaged across all participants for EII and MI approaches.

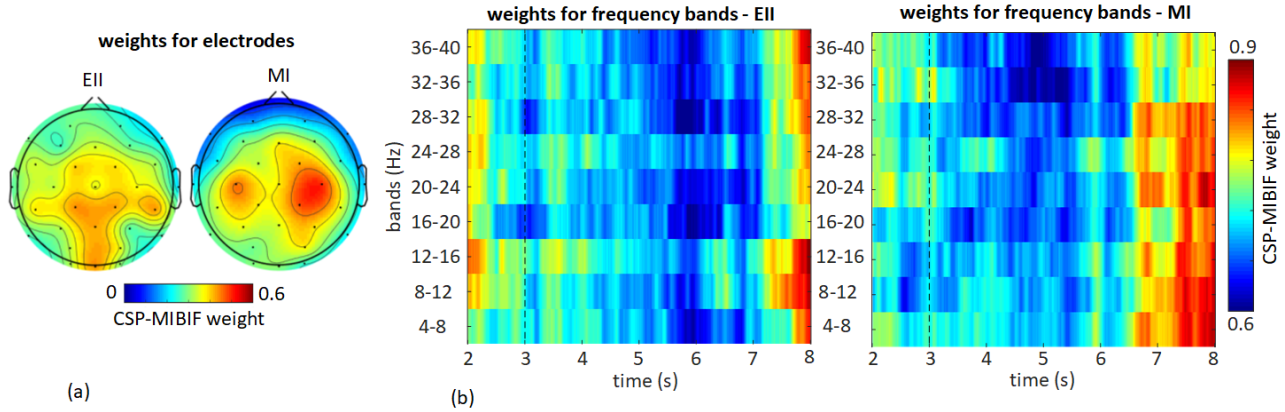

Fig. 7. *The CSP-mutual information best individual feature (MBIF) weights (a) for each electrodes at the time of peak cross-validation accuracy averaged across participants and (b) time-course weights for frequency bands for the two BCI approaches (EII and MI); the vertical dotted line in (b) indicate the task's cue, at 3s. The time-course in (b) starts at 2s because the analysis was based on a 2s sliding window.*

that the performance of the two imagery types are significantly different, $F(1, 11) = 8.45$, $p < 0.05$, $\eta^2 = 0.44$; the performance with MI is significantly higher than performance with EII ($p < 0.05$, Bonferroni corrected). The ANOVA results also show that there is a significant difference of performance across the BCI frameworks, $F(4, 44) = 13.79$, $p < 0.001$, $\eta^2 = 0.56$. Overall for MI and EII, pairwise comparisons (Bonferroni corrected) shows that the performance with ASYM framework is significantly lower than the performance with NTSPP framework ($p < 0.05$), significantly lower than the performance with NTSPP-FBCSP framework ($p < 0.05$), significantly lower than the performance with COMB ($p < 0.001$), but not significantly different from the performance with FBCSP-framework ($p = 0.21$). The performance achieved with COMB framework is significantly higher than the performance with FBCSP framework ($p < 0.05$), but not different from performance achieved with NTSPP framework ($p = 1$) and not significantly different from the performance with NTSPP-FBCSP framework ($p = 0.11$). The performance with NTSPP framework was not significantly higher than the performance with FBCSP framework ($p = 0.25$)

Further pairwise comparisons in EII classification accuracies show that the performances of ASYM and FBCSP are each significantly lower than the performance of COMB ($p < 0.001$ and $p < 0.005$ respectively), but they are not significant different from the performances of the NTSPP ($p = 1$ and $p = 1$ respectively) and NTSPP-FBCSP ($p = 1$ and $p < 0.769$ respectively). With the EII strategy, the performance of COMB is not significantly higher than the performance of NTSPP ($p = 0.167$) and NTSPP-FBCSP ($p = 0.087$). For MI, the performance of ASYM is significantly lower than the performance achieved with NTSPP ($p < 0.05$), NTSPP-FBCSP ($p < 0.05$), and COMB ($p < 0.005$). Also the NTSPP did not perform significantly better than FBCSP ($p = 0.479$), NTSPP-FBCSP ($p = 1$), and COMB ($p = 1$) for MI strategy.

B. The Weighted Contribution of Different Electrodes and Frequency Bands

The contribution of different scalp areas (electrodes) toward MI and EII performance across participants is shown in Fig. *7(a)*. The electrodes contribution is shown as topographic map of CSP-MIBIF weights averaged across participants at the time of peak cross-validation accuracy. For MI, the right and left central electrodes (C3, C4, and CP2) are most important. On the other hand, the most weighted electrodes for EII are located in central-parietal area (CP1, CP2, Cp6), parietal (PZ), and occipital (OZ). Comparing the color map of the most relevant electrodes, the weighting in the case of MI is greater than those in EII which suggests higher consistency in electrodes from MI.

The weighted contribution of different frequency bands is shown Fig. *7(b)*. The frequency band contributions are presented as a time-course of CSP-MIBIF weights for each frequency band averaged across participants. The frequency bands covering 4-16Hz and 20-40Hz are weighted most for EII, with most prominent weighting in the last 0.5s of the task execution across participants. For the MI case, most of the weight is distributed from frequency bands covering 4 to 32Hz, with most prominent weighting in the last 1.5s of the task execution.

C. Participants' Subjective Responses

In 88.24% of the sessions, the participants reported that MI was the favorite approach versus 11.76% for EII. Also, participants reported that MI is the approach in which they perceived most control in 89.71% of the sessions versus 10.29% for EII.

D. Correlation between Pre-run EEG and Performance

Participants showed different associations between their performance and pre-run band power ratios. The results for the participants with significant correlation are shown in the TABLE 1, for all the significant correlation cases, $r^2 > 0.8$. For both EII and MI, we found no general trend across participants for correlation between their pre-run power ratios and performance. There are a few instances of opposite correlation for MI and EII cases, but not consistent across participants. The participant *ay* showed positive correlation of pre-run middle-line occipital theta with MI but a negative correlation with EII performance. On the other hand, the participant *kp* showed positive correlation of left parietal theta with EII but negative correlation with MI performance. Participant *kp* also showed positive correlation of middle-line parietal alpha with MI but

TABLE 1
CORRELATION RESULTS OF PRE-RUN EEG AND CLASSIFICATION ACCURACY ACROSS SESSIONS

Scalp locations\BCI	EII	MI	EII	MI	EII	MI	EII	MI	EII	MI	EII	MI	EII	MI
Frontal-middle											*a			
Temporal-left							*t			*g				
Central-right		*g	*a							*b				
Parietal-left		**-t *b *g					*t *-b	*-t			**t			
Parietal-middle						*a	**-a	*a						
Parietal-right	*b						*t							
Occipital-left	*a	*g				*-b	*-a *g			*t				
Occipital-middle					*-t *b	**t								
Occipital-right						***-g								*-t
Participants IDs	ak		bt		ay		kp		ar		tf		pw	

a = alpha, b = beta, g= gamma, t= theta band **: p < 0.05, **: p < 0.01, ***: p < 0.005* *-: negative correlation*

negative correlation with EII performance.

IV. DISCUSSION

A. Performance of EII versus MI

The main aim of this study was to compare performance of emotion-inducing imagery versus motor imagery BCI strategies across multiple recording sessions involving online visual feedback. A preliminary single-session based study previously reported comparable performance between EII and MI [23], however the results presented here show that performance with MI is significantly higher than EII performance. Nine out of the twelve participants performed above chance level CA in all their EII runs, but only 4 participants (*bt*, *ay*, *jl* and *pw*) achieved mean online *CA ≥ 70%*, usually considered as acceptable performance for BCI setup used in this study [39]. Some of the participants (i.e. *ak*, *kp*, *ar*, and *ut*) showed acceptable MI performance but poor EII performance. Some of the participants (i.e. *ay* and *tf*) performed better with EII than with MI.

The results from re-simulating online runs using various frameworks show that MI outperforms EII. These results also show that among BCIs tested (NTSPP, FBCSP, ASYM, NTSPP-FBCSP, and COMB), the best performing BCI setup is NTSPP for MI which is also the setup used during the online runs. To our knowledge, this is the first time the NTSPP is compared to FBCSP in a multi-session MI. For EII, the combination of all the frameworks led to best performing framework but not significantly different from the NTSPP-based setup used in the actual EII feedback runs.

Four participants achieved acceptable performance (above 70% for CA) with EII, and one of them (i.e. *ay*) performed better than with MI, as shown in Fig. 5. This result suggests that these four participants, especially '*ay*', may use EII as an alternative imagery approach for BCI. On the other hand, eight participants achieved acceptable accuracy with MI, and six of them performed better than with EII. These results suggest a hybrid MI-EII BCI may suit some BCI users. Since in our study the tasks used in EII were limited to imagining (or recalling) sad and happy events, future work should consider widening the range of imagery tasks selection to include emotional faces, scenes, objects, pictures, words, and sounds. With a wide range of imagery tasks, BCI users, especially those with poor MI performance, are likely to find suitable imagery tasks that may lead to improved performance in classification accuracy and comfort of use.

Apart from the superior CA performance of MI over EII, participants preferred MI to EII as a BCI control approach. This preference is likely influenced by difficulty experienced in accessing repeatedly similar events in memory in a short time-period. Furthermore, recalling emotional events may trigger a series of memories, which make it hard to focus on the targeted event. This issue could be alleviated by asking participants to recall attributes or objects [40] which can be associated with some emotional events or with emotional states, e.g., odors, as in [11].

B. Frameworks Performance Comparison

In this study we found that the performance of MI was consistently higher than the performance of EII across NTSPP, FBCP, NTSPP-FBCSP and COMB frameworks. The NTSPP framework averaged the highest performance across participants for MI but not in the case of EII. The average performance of EII is higher with a signal processing framework that combines the NTSPP, FBCSP and ASYM frameworks (COMB). The observed high EII performance could be due to the COMB framework accessing several different features associated with the complexity of emotion

related cortical networks and neural circuitry [41]. Having different setups combined into the COMB framework and each setup individually discriminating the classes by using different features, this allows extraction of features associated to a variety of neural processes of interest. The resulting decision for COMB is likely to consider relevant features associated to EII processes in the brain. Future EII studies should also investigate the effect of training a BCI using only data selected based on relationship across trials, e.g., correlation-based time window selection as in [42], instead of a fixed time window across trials used in current study.

C. Pre BCI-use and BCI Performance Relationship

Through frequency content analysis of brain activity before feedback run, we identified different scalp areas and frequency bands correlating with performance in EII and MI for most participants, but we did not observe any common topographic correlation pattern across these participants for either EII or MI. Previously, Ahn and colleagues [43], from recordings runs scheduled on the same day, have reported significant correlation of gamma power ratio during resting state magnetoencephalogram (MEG) and motor imagery classification performance achieved with EEG (simultaneously recorded with MEG). In our study, only three of participants (i.e. *ak*, *ay*, *ar*), with good performance in MI, showed correlation between gamma power ratio and MI performance across sessions, and for one of these participants, *ay*, this was a negative correlation. In our study, the correlation analysis involves several sessions scheduled at different days for each individual participant, which may be the reason we did not observe the previously reported gamma correlation pattern throughout the participants with good performance in MI. The knowledge of how pre-run power ratios are associated with a given participant's performance may enable monitoring and identification of the participant's optimal state for good performance and/or adaption and selection of BCI parameters. Future effort should focus on identifying activity or tasks that increase or decrease targeted spectral content in neural signals during the pre-BCI use period and the influence of time of day and other behavior related variables.

D. Frequency Bands and Electrodes Contribution to MI and EII Performance

The results for relevant electrodes during task execution show that the electrodes mounted on sensorimotor cortex are most important for MI across participants. This finding is in line with the claim that a MI task activate sensorimotor cortex [44]. For EII, the electrodes weighted the heaviest are in central-parietal, parietal and occipital areas. The weights for relevant electrodes across participants are less in than EII case than in MI. The light weights of electrodes after averaging across participants, in EII case, suggest that spatial patterns and frequencies engaged by EII tasks are less consistent across participants. EII task may activate several areas of the brain including the frontal, temporal and visual areas [45] depending on the imagery vividness, and this might be the reason we observe non-negligible weights for frontal electrodes with EII in Fig. *7(a)*.

The time-course of weights for frequency bands relevancy across participant shows that after the cue (3s), the weights decrease (mostly visible from 4 to 6s, in *Fig. 7(b)*), and then increase for both EII and MI. The high weights indicate most relevant frequency bands across participants at a given time during the task execution. The high weights are mainly observed during the last 0.5s in the task execution for EII whereas for MI this is observed during the last 1.5s. The high weights observed at the end of the task execution for EII case suggests that the EII tasks classification performance may benefit from increasing the task execution period. The high weights being sustained for short period in EII case may be due to variation of imageries used in EII tasks across participants. Future studies on EII should consider minimizing variation in imageries, by using imagery of specific objects or attributes commonly associated with emotional states (e.g., baby's laugh, which could be associated with happy event) consistently across participants as stimuli.

The game paradigm was used in our study to maintain motivation and curiosity in the participants. However, the feedback involved a game character moving toward left or right, which create intuitive correspondence with imagery for left or right hand movement compared to imagery for happy or sad event. A more neutral feedback like Graz BCI [46] may allow to accurately assess the differences in spatial and frequency patterns associated with EII and MI.

V. Conclusion

The comparison between motor imagery and emotion-inducing imagery as BCI control strategies across several recording sessions with online feedback control showed that MI outperforms EII. The MI performed best with NTSPP among the signal processing frameworks considered which included FBCSP, a state-of-the-art framework for MI based BCI. We found that EII performance benefits from fusing various EEG features, and this finding should be exploited for future EII studies.

EII may offer a viable alternative in some cases for subjects who cannot control a motor imagery BCI, but further investigation is necessary to identify effective EII tasks that might be easy to execute in a BCI paradigm and potentially be combined with MI tasks to create a multi-class and/or hybrid BCI.

Acknowledgment

The authors wish to thank the volunteers who participated in this study.

References

[1] D. Coyle, J. Garcia, A. R. Satti, and T. M. McGinnity, "EEG-based continuous control of a game using a 3 channel motor imagery BCI: BCI game," in *2011 IEEE Symposium on Computational Intelligence, Cognitive Algorithms, Mind, and Brain (CCMB)*, 2011, pp. 1–7.

[2] M. Ahn, M. Lee, J. Choi, and S. C. han Jun, "A review of brain-computer interface games and an opinion survey from researchers, developers and users," *Sensors (Basel).*, vol. 14, no. 8, pp. 14601–14633, 2014.

[3] Y. Yu *et al.*, "Toward brain-actuated car applications: Self-paced control with a motor imagery-based brain-computer interface," *Comput. Biol. Med.*, vol. 77, pp. 148–155, 2016.

[4] G. Prasad, P. Herman, D. Coyle, S. McDonough, and J. Crosbie, "Applying a brain-computer interface to support motor imagery practice in people with stroke for upper limb recovery: a feasibility study," *J. Neuroeng. Rehabil.*, vol. 7, no. 1, p. 60, Dec. 2010.

[5] J. R. Wolpaw and E. Winter Wolpaw, "Brain–Computer Interfaces: Something New under the Sun," in *Brain–Computer InterfacesPrinciples and Practice*, vol. 6, no. 38, Oxford University Press, 2012, pp. 3–12.

[6] B. Blankertz *et al.*, "Neurophysiological predictor of SMR-based BCI performance," *Neuroimage*, vol. 51, no. 4, pp. 1303–1309, 2010.

[7] M. Ahn, H. Cho, S. Ahn, and S. C. Jun, "High Theta and Low Alpha Powers May Be Indicative of BCI-Illiteracy in Motor Imagery," *PLoS One*, vol. 8, no. 11, p. e80886, Nov. 2013.

[8] M. C. Thompson, "Critiquing the Concept of BCI Illiteracy," *Sci. Eng. Ethics*, vol. 25, no. 4, pp. 1217–1233, Aug. 2019.

[9] S. Makeig, G. Leslie, T. Mullen, D. Sarma, N. Bigdely-Shamlo, and C. Kothe, "First Demonstration of a Musical Emotion BCI," in *Affective Computing and Intelligent Interaction*, 2011, pp. 487–496.

[10] G. Chanel, J. J. M. Kierkels, M. Soleymani, and T. Pun, "Short-term emotion assessment in a recall paradigm," *Int. J. Hum. Comput. Stud.*, vol. 67, no. 8, pp. 607–627, 2009.

[11] D. Iacoviello, A. Petracca, M. Spezialetti, and G. Placidi, "A real-time classification algorithm for EEG-based BCI driven by self-induced emotions," *Comput. Methods Programs Biomed.*, vol. 122, no. 3, pp. 293–303, 2015.

[12] R. Sitaram, S. Lee, S. Ruiz, M. Rana, R. Veit, and N. Birbaumer, "Real-time support vector classification and feedback of multiple emotional brain states," *Neuroimage*, vol. 56, no. 2, pp. 753–765, May 2011.

[13] S. D. Power, A. Kushki, and T. Chau, "Towards a system-paced near-infrared spectroscopy brain–computer interface: differentiating prefrontal activity due to mental arithmetic and mental singing from the no-control state," *J. Neural Eng.*, vol. 8, no. 6, p. 066004, Oct. 2011.

[14] M. Stangl, G. Bauernfeind, J. Kurzmann, R. Scherer, and C. Neuper, "A haemodynamic brain-computer interface based on real-time classification of near infrared spectroscopy signals during motor imagery and mental arithmetic," *J. Near Infrared Spectrosc.*, vol. 21, no. 3, pp. 157–171, 2013.

[15] E. V. C. Friedrich, R. Scherer, and C. Neuper, "The effect of distinct mental strategies on classification performance for brain-computer interfaces," *Int. J. Psychophysiol.*, vol. 84, no. 1, pp. 86–94, 2012.

[16] E. Curran *et al.*, "Cognitive Tasks for Driving a Brain-Computer Interfacing System: A Pilot Study," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 12, no. 1, pp. 48–54, 2004.

[17] R. J. Davidson, P. Ekman, C. D. Saron, J. A. Senulis, and W. V Friesen, "Approach-withdrawal and cerebral asymmetry: emotional expression and brain physiology I," *Journal of personality and social psychology*, vol. 58, no. 2. pp. 330–341, 1990.

[18] J. J. B. Allen, E. Harmon-Jones, and J. H. Cavender, "Manipulation of frontal EEG asymmetry through biofeedback alters self-reported emotional responses and facial EMG," *Psychophysiology*, vol. 38, no. 4, pp. 685–693, 2001.

[19] T. Canli, "Hemispheric Asymmetry in the Experience of Emotion: A Perspective from Functional Imaging," *The Neuroscientist*, vol. 5, no. 4. pp. 201–207, 1999.

[20] C. A. Kothe, S. Makeig, and J. A. Onton, "Emotion Recognition from EEG during Self-Paced Emotional Imagery," in *2013 Humaine Association Conference on Affective Computing and Intelligent Interaction*, 2013, pp. 855–858.

[21] E. E. Fetz, "Volitional control of neural activity: implications for brain-computer interfaces," *J. Physiol.*, vol. 579, no. 3, pp. 571–579, 2007.

[22] A. D. Bigirimana, N. Siddique, and D. Coyle, "Emotion Imagery BCI," in *Proceedings of the 6th International Brain-Computer Interface Meeting*, 2016, pp. 125–125.

[23] A. D. Bigirimana, N. Siddique, and D. Coyle, "Brain-Computer Interfacing with Emotion-Inducing Imagery: A Pilot Study," in *7th Graz Brain-Computer Interface Conference (GBCIC)*, 2017, pp. 26–31.

[24] D. Coyle, G. Prasad, and T. M. McGinnity, "A time-series prediction approach for feature extraction in a brain-computer interface.," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 13, no. 4, pp. 461–467, 2005.

[25] D. Coyle, "Neural network based auto association and time-series prediction for biosignal processing in brain-computer interfaces," *IEEE Comput. Intell. Mag.*, vol. 4, no. 4, pp. 47–59, 2009.

[26] K. K. Ang, Z. Y. Chin, C. Wang, C. Guan, and H. Zhang, "Filter Bank Common Spatial Pattern Algorithm on BCI Competition IV Datasets 2a and 2b," *Front. Neurosci.*, vol. 6, no. MAR, pp. 1–9, 2012.

[27] S. M. Alarcao and M. J. Fonseca, "Emotions Recognition Using EEG Signals: A Survey," *IEEE Trans. Affect. Comput.*, vol. 3045, no. c, pp. 1–20, 2017.

[28] G. Pfurtscheller *et al.*, "Current trends in Graz Brain-Computer Interface (BCI) research," *IEEE Trans. Rehabil. Eng.*, vol. 8, no. 2, pp. 216–219, 2000.

[29] F. Lotte and Cuntai Guan, "Regularizing Common Spatial Patterns to Improve BCI Designs: Unified Theory and New Algorithms," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 2, pp. 355–362, Feb. 2011.

[30] D. Coyle, G. Prasad, and T. M. McGinnity, "Faster self-organizing fuzzy neural network training and a hyperparameter analysis for a brain-computer interface," *IEEE Trans. Syst. Man, Cybern. Part B Cybern.*, vol. 39, no. 6, pp. 1458–1471, 2009.

[31] G. Pfurtscheller *et al.*, "Separability of EEG signals recorded during right and left motor imagery using adaptive autorregresive parameters," *IEEE Trans. Rehabil. Eng.*, vol. 6, no. 3, pp. 316–324, 1998.

[32] M.-K. Kim, M. Kim, E. Oh, and S.-P. Kim, "A Review on the Computational Methods for Emotional State Estimation from the Human EEG," *Comput. Math. Methods Med.*, vol. 2013, pp. 1–13, 2013.

[33] A. Schlögl, C. Neuper, and G. Pfurtscheller, "Estimating the Mutual Information of an EEG-based Brain-Computer Interface," *Biomed. Tech. Eng.*, vol. 47, no. 1–2, pp. 3–8, 2002.

[34] Kai Keng Ang, Chai Quek, K. K. Ang, and C. Quek, "Rough set-based neuro-fuzzy system," *2006 IEEE Int. Jt. Conf. Neural Netw. Proc.*, pp. 742–749, 2006.

[35] A. Delorme and S. Makeig, "EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis," *J. Neurosci. Methods*, vol. 134, no. 1, pp. 9–21, 2004.

[36] N. P. Castellanos and V. a. Makarov, "Recovering EEG brain signals: Artifact suppression with wavelet enhanced independent component analysis," *J. Neurosci. Methods*, vol. 158, pp. 300–312, 2006.

[37] A. D. Bigirimana, N. Siddique, and D. Coyle, "A hybrid ICA-wavelet transform for automated artefact removal in EEG-based emotion recognition," in *2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2016, pp. 004429–004434.

[38] G. R. Müller-putz, R. Scherer, C. Brunner, R. Leeb, and G. Pfurtscheller, "Better than random ? A closer look on BCI results," *Int. Jouranl Bioelectromagn.*, vol. 10, no. 1, pp. 52–55, 2008.

[39] C. Vidaurre and B. Blankertz, "Towards a cure for BCI illiteracy," *Brain Topogr.*, vol. 23, no. 2, pp. 194–198, 2010.

[40] A. Korik *et al.*, "Primitive shape imagery classification from electroencephalography," *7th Int. BCI Meet.*, no. August, pp. 2–3, 2018.

[41] T. Dalgleish, "The emotional brain.," *Nat. Rev. Neurosci.*, vol. 5, no. 7, pp. 582–589, 2004.

[42] J. Feng *et al.*, "Towards correlation-based time window selection method for motor imagery BCIs," *Neural Networks*, vol. 102, pp. 87–95, 2018.

[43] M. Ahn *et al.*, "Gamma band activity associated with BCI performance: simultaneous MEG/EEG study," *Front. Hum. Neurosci.*, vol. 7, no. December, pp. 1–10, 2013.

[44] G. Pfurtscheller and C. Neuper, "Motor imagery activates primary sensorimotor area in humans," *Neurosci. Lett.*, vol. 239, no. 2–3, pp. 65–68, 1997.

[45] D. R. Addis, A. T. Wong, and D. L. Schacter, "Remembering the past and imagining the future: Common and distinct neural substrates during event construction and elaboration," *Neuropsychologia*, vol. 45, no. 7, pp. 1363–1377, 2007.

[46] G. Pfurtscheller, C. Neuper, and N. Birbaumer, "Human Brain-Computer Interface," in *Motor Cortex in virtual Movemen*, A distribu., Wahington, DC: CRC Press, 2005, pp. 367–401.